

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO

Implementação de ferramenta para geração automática de documentação estruturada

Cátia Cruz

DISSERTAÇÃO



Mestrado Integrado em Engenharia Informática e Computação

Orientador: Nuno Flores (PhD)

Orientador na empresa: Pedro Silva

29 de Julho de 2015

Implementação de ferramenta para geração automática de documentação estruturada

Cátia Cruz

Mestrado Integrado em Engenharia Informática e Computação

Resumo

Ao longo dos últimos anos, cada vez mais empresas têm investido os seus recursos em sistemas de gestão documental, em busca da melhoria dos seus processos de trabalho para, deste modo, alcançarem produtividades que as permitam crescer no mercado de trabalho. É com esta premissa que a Finantech decidiu migrar toda a informação que tinha guardada em documentos words guardados em pastas, para uma wiki integrada no SharePoint, aperfeiçoando assim, o seu sistema de gestão documental. Surge então a necessidade de criar uma ferramenta (completamente integrada no SharePoint) que permita gerar documentos word, com os conteúdos agora guardados na wiki, para quando é necessário produzir documentação para impressão.

Nesta dissertação encontra-se descrito todo o processo de criação desta ferramenta - a Wiki2Doc - que permite retirar automaticamente os conteúdos da wiki da empresa (integrada no SharePoint), possibilitando a geração da documentação estruturada desejada. São explicados alguns conceitos e é feito um estudo de algumas aplicações existentes no mercado, relacionadas com o trabalho. É ainda apresentado esquematicamente o funcionamento, as tecnologias adotadas e os métodos de implementação e integração do protótipo na plataforma SharePoint. Por último, são introduzidas algumas limitações da ferramenta, juntamente com as conclusões e trabalho futuro a desenvolver pela Finantech.

Abstract

Over the last years, more companies have invested their resources in document management systems in order to improve their work procedures, which will allow them to achieve high productivity standards and grow in the market. Based on this premise Finantech decided to transfer all the contents and information stored in word files to a wiki integrated in the SharePoint platform. Because of this decision Finantech felt the need to build a tool that creates word documents, with the information stored in the wiki, for matters of producing printing documents to deliver to their clients. The purpose of this project is to develop a prototype, fully integrated in SharePoint that creates manuals, technical reports and other types of documents from templates designed by Finantech, with the information stored in the wiki.

This thesis describes the designing and making process of this tool - Wiki2Doc - that automatically withdraws information from the wiki(integrated in SharePoint), and generates the required documentation. Some key concepts are also explained and it's held a study about the applications already available in the market. Furthermore this paper describes the technologies and procedures that were used to develop the tool and incorporate it in the SharePoint platform. Lastly, are described some of the tool's limitations along with the conclusions and the future work.

Agradecimentos

A entrega desta dissertação marca o final de uma etapa na minha vida e o início de outra ainda mais importante. Foram muitas as pessoas que me acompanharam ao longo deste período, a quem gostava de agradecer.

Em primeiro lugar, gostaria de expressar a minha gratidão ao Eng. Nuno Flores da FEUP por ter aceitado ser meu orientador, e por todo o apoio e incentivo que me deu durante estes 4 meses.

Gostava também de agradecer à Finantech por me ter proporcionado uma experiência fantástica e a todas as pessoas do Departamento de Integration and Data Analysis (IDA), em especial, ao Pedro Silva e à Patrícia Nogueira por toda a paciência que tiveram e por toda a ajuda e apoio que me deram.

Aos meus amigos, Bárbara Correia, Damien Rosa, Maria João Gabriel, Maling Wang e Rita Mendes, pela sua amizade e lealdade, pelo carinho constante que me deram e dão e pelo apoio e força que me deram durante a realização desta dissertação.

Ao meu pai, João Cruz, que me ajudou muito durante todo o meu percurso no MIEIC.

Às minhas irmãs Ana Sofia e Joana por serem minhas amigas e confidentes e ao Filipe Machado por ser como um pai para mim.

Finalmente, um agradecimento muito especial à minha mãe, Anabela Rodrigues, à minha psicóloga, Dra. Paula Morais e ao meu namorado, André Gonçalves pela paciência infinita e pelo apoio constante e incansável. Foram os três pilares mais importantes da minha vida neste último ano.

Sem cada uma destas pessoas, teria sido muito mais árdua a chegada ao final desta jornada. O meu sincero obrigado a todos vós.

Cátia Cruz

“Try not to become a man of success, but rather try to become a man of value.”

Albert Einstein

Conteúdo

| | | |
|----------|---|-----------|
| 1 | Introdução | 1 |
| 1.1 | Contexto e Enquadramento | 1 |
| 1.2 | Objetivos | 2 |
| 1.3 | Estrutura da Dissertação | 3 |
| 2 | Trabalho relacionado | 5 |
| 2.1 | Conceitos Base | 5 |
| 2.1.1 | Gestão documental | 5 |
| 2.1.2 | Geração automática de documentos | 6 |
| 2.1.3 | Wiki | 8 |
| 2.1.4 | HTML | 8 |
| 2.1.5 | DOCX | 8 |
| 2.1.6 | XML | 8 |
| 2.1.7 | SharePoint | 8 |
| 2.2 | Conversores HTML para DOCX | 10 |
| 2.2.1 | docusuit_ | 10 |
| 2.2.2 | PHPWord | 10 |
| 2.2.3 | Convert Doc | 11 |
| 2.2.4 | Conclusões | 11 |
| 3 | Wiki2Doc | 13 |
| 3.1 | Motivação | 13 |
| 3.2 | Requisitos e Restrições da Ferramenta | 13 |
| 3.2.1 | Requisitos Funcionais | 13 |
| 3.2.2 | Requisitos Não-Funcionais | 14 |
| 3.3 | Arquitetura | 14 |
| 3.3.1 | Visão Geral | 14 |
| 3.3.2 | Arquitetura Lógica | 15 |
| 3.3.3 | Casos de Uso | 16 |
| 3.4 | Tecnologias usadas | 16 |
| 3.4.1 | Open XML SDK 2.5 | 16 |
| 3.4.2 | HtmlAgilityPack | 18 |
| 3.4.3 | SharePoint | 18 |
| 3.5 | Implementação | 19 |
| 3.5.1 | Conversor de Html para Word | 20 |
| 3.5.2 | Integração do conversor no SharePoint | 20 |
| 3.6 | Desenvolvimento da Interface | 24 |
| 3.7 | Limitações da Ferramenta | 26 |

CONTEÚDO

| | | |
|----------|---|-----------|
| 3.7.1 | Interface na edição de Templates | 26 |
| 3.7.2 | Controlo de Versões | 26 |
| 3.7.3 | Estrutura das secções dos Templates | 27 |
| 3.7.4 | Dificuldades Sentidas | 27 |
| 3.8 | Conclusões Retiradas | 27 |
| 4 | Conclusões e Trabalho Futuro | 29 |
| 4.1 | Conclusões | 29 |
| 4.2 | Trabalho Futuro | 30 |
| | Referências | 31 |
| A | Templates e Manuais Gerados | 33 |

Lista de Figuras

| | | |
|------|--|----|
| 1.1 | Mapa Gantt do Projeto | 2 |
| 2.1 | Arquitetura Principal do SharePoint | 9 |
| 3.1 | Visão Geral | 15 |
| 3.2 | Arquitetura Lógica do Sistema | 15 |
| 3.3 | Casos de Uso do Utilizador | 16 |
| 3.4 | Exemplo da estrutura básica de um documento WordprocessingML | 17 |
| 3.5 | Elementos de um documento WordprocessingML | 17 |
| 3.6 | Representação HTMLDocument | 18 |
| 3.7 | Estrutura do Site da Finantech | 19 |
| 3.8 | Representação do elemento XElement do tipo Altchunk | 20 |
| 3.9 | Índice da wiki integrada no SharePoint (1º Nivel) | 21 |
| 3.10 | Índice da wiki integrada no SharePoint (2º Nivel - Ferramenta SOAPP) | 22 |
| 3.11 | Exemplo de página wiki integrada no SharePoint (“O Meu Top 20”) | 22 |
| 3.12 | Template com secção correspondente à página “O Meu Top 20” | 23 |
| 3.13 | Fluxo de passagem da informação | 24 |
| 3.14 | Página inicial do SharePoint | 24 |
| 3.15 | Página Inicial da Wiki integrada no SharePoint | 25 |
| 3.16 | Menu “Matriz de Manuais” | 25 |
| 3.17 | Menu “Gerar Manual” | 26 |
| 3.18 | Menu “Manuais Gerados” | 26 |
| A.1 | Template de Exemplo 1 (Pág.1) | 34 |
| A.2 | Template de Exemplo 1 (Pág.2) | 35 |
| A.3 | Template de Exemplo 1 (Pág.3) | 36 |
| A.4 | Template de Exemplo 1 (Pág.4) | 37 |
| A.5 | Manual de Exemplo 1 (Pág.1) | 38 |
| A.6 | Manual de Exemplo 1 (Pág.2) | 39 |
| A.7 | Manual de Exemplo 1 (Pág.3) | 40 |
| A.8 | Manual de Exemplo 1 (Pág.4) | 41 |
| A.9 | Manual de Exemplo 1 (Pág.5) | 42 |
| A.10 | Template de Exemplo 2 (Pág.1) | 43 |
| A.11 | Template de Exemplo 2 (Pág.2) | 44 |
| A.12 | Template de Exemplo 2 (Pág.3) | 45 |
| A.13 | Template de Exemplo 2 (Pág.4) | 46 |
| A.14 | Manual de Exemplo 2 (Pág.1) | 47 |
| A.15 | Manual de Exemplo 2 (Pág.2) | 48 |
| A.16 | Manual de Exemplo 2 (Pág.3) | 49 |

LISTA DE FIGURAS

| | |
|----------------------------------|----|
| A.17 Manual de Exemplo 2 (Pág.4) | 50 |
|----------------------------------|----|

Abreviaturas e Símbolos

| | |
|------|-------------------------------|
| MTF | Multilateral Trading Facility |
| IDA | Integration and DataAnalysis |
| HTML | Hypertext Markup Language |
| XML | EXtensible Markup Languag |
| IIS | Internet Information Service |
| URL | Uniform Resource Locator |
| BD | Base de Dados |
| PDF | Portable Document Format |
| RTF | Rich Text Format |
| ODT | Open Document Text |
| PHP | PHP: Hypertext Preprocessor |
| CSV | Comma Separated Values format |
| TXT | Text File |
| SQL | Structured Query Language |
| SDK | Software Development Kit |

Capítulo 1

Introdução

Este documento ilustra todo o desenvolvimento e implementação de uma ferramenta de geração automática de documentos, integrada no SharePoint 2013, o qual se denominou de Wiki2Doc.

Relativamente ao conteúdo deste primeiro capítulo, após a apresentação do tema da dissertação, é referido o contexto e enquadramento no qual o projeto foi desenvolvido. Segue-se uma breve descrição dos objetivos do projeto e, por último, é mencionada a restante estrutura do documento.

1.1 Contexto e Enquadramento

Numa busca constante pelo aumento da qualidade no local de trabalho, as empresas sentem, cada vez mais, a necessidade de ter a sua documentação organizada/estruturada e preservada de forma a facilitar a administração da mesma. A gestão de documentos é essencial ao bom desempenho de qualquer empresa e um reflexo direto da sua organização interna. A existência de grandes volumes de informação (tanto oriunda do exterior como interna) potencia a perda de documentos, dificulta o acesso à informação (e à sua localização física) e prejudica ainda o controlo dos fluxos de informação.

A Finantech é uma empresa que se dedica ao desenvolvimento e manutenção de aplicações de suporte às funções de “BackOffice”, “MiddleOffice” e “FrontOffice” para a Banca e outros Intermediários Financeiros. Possui soluções dedicadas a sistemas de negociação online de Brokers, custódia e liquidação para Bolsas e MTFs, gestão de patrimónios, redes de balcões e lojas financeiras. A gestão documental da Finantech baseava-se em documentos Word guardados em pastas, o que causava diversos problemas na consulta da informação por parte dos funcionários da empresa. Com o passar dos anos, o aumento do volume de negócios da empresa conduziu a um elevado acréscimo na quantidade de informação processada e documentada, suscitando assim a necessidade da empresa ter a sua documentação mais organizada e estruturada. Desta forma, de modo a facilitar o acesso à informação por parte dos seus funcionários, a empresa decidiu adotar o

Introdução

software SharePoint como plataforma de gestão documental e criar uma wiki (totalmente integrada no Sharepoint) para onde se está a migrar toda a informação contida nos documentos word, permitindo que todo o conhecimento da empresa esteja organizado de forma transversal, sistemática e eficiente.

A plataforma de gestão documental está a ser desenvolvida pela área de Integration and Data Analysis (IDA) da empresa desde o início de 2015. Tendo em conta a curta duração deste projeto, existem ainda muitas carências no sistema a serem solucionadas. Uma das principais dificuldades com que os funcionários se deparavam era o facto de terem de estar constantemente a recorrer à wiki para criar os documentos (copiando manualmente o conteúdo pretendido, correndo o risco de perder informação este processo). Assim sendo, em março do ano corrente, foi aprovado pela empresa o desenvolvimento de uma ferramenta que transporta automaticamente os conteúdos da wiki, integrada no SharePoint, para documentos Word, utilizando templates prédefinidos pela Finantech.

Abaixo encontra-se um gráfico temporal, representativo das diferentes fases do trabalho.

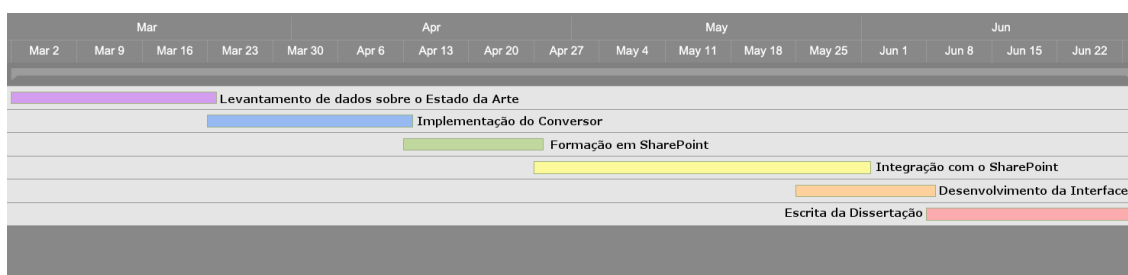


Figura 1.1: Mapa Gantt do Projeto

1.2 Objetivos

O propósito deste trabalho consiste em desenvolver uma ferramenta que transporte automaticamente os conteúdos do arquivo digital centralizado (as páginas da wiki integrada no SharePoint) para documentos word (docx), utilizando templates predefinidos pela Finantech, gerando assim os documentos desejados (manuais internos e para clientes, documentação técnica e de análise). Através desta ferramenta, pretende-se que os funcionários possam escolher intuitivamente o conteúdo da wiki a colocar em cada secção do documento a gerar e que o mesmo seja guardado automaticamente no SharePoint. Desta forma, prevê-se não só facilitar e melhorar intuitivamente toda a organização e estruturação dos conteúdos documentais, mas também potenciar a vertente de negócio da Finantech (através da criação e manutenção de documentação final estruturada e organizada para clientes existentes e prospetivos).

1.3 Estrutura da Dissertação

Para além da introdução, este documento é composto por mais 3 capítulos. No capítulo 2, encontram-se descritos alguns conceitos base necessários para o desenvolvimento do projeto e são referidas algumas das ferramentas existentes no mercado relacionadas com este trabalho, nomeadamente, conversores de html para documentos Word (docx).

O Capítulo 3 diz respeito às especificações do projeto. É apresentada a motivação do projeto, seguida da descrição pormenorizada da arquitetura do sistema (incluindo uma análise da visão geral, da arquitetura lógica e dos casos de uso). Segue-se um estudo das várias tecnologias utilizadas, da implementação do projeto e do desenvolvimento da interface. Finalmente são referidas limitações quanto à ferramenta Wiki2Doc.

As conclusões retiradas e as referências para trabalho futuro são referidas no capítulo 4, o último capítulo deste documento.

Introdução

Capítulo 2

Trabalho relacionado

Neste capítulo, encontram-se descritos alguns conceitos que servem de suporte a este trabalho (necessários para a compreensão da ferramenta desenvolvida) e são apresentadas algumas aplicações e bibliotecas relacionadas com o projeto, nomeadamente alguns conversores de html para ficheiros word.

2.1 Conceitos Base

De forma a melhor assimilar todo o processo de implementação da ferramenta desenvolvida, foram explorados alguns conceitos base que facilitam a compreensão das várias fases do projeto.

2.1.1 Gestão documental

A implementação de um sistema de gestão documental é um fator decisivo para o negócio de qualquer empresa, na medida em que este permite gerir eficazmente toda a informação processada. São vastos os benefícios da implementação de soluções de gestão documental, variando desde a redução do tempo gasto (no processamento manual de documentos) até à precisão dos materiais acessíveis a todo o tipo de utilizadores destes sistemas.

Alguns dos benefícios mais comuns que valorizam as empresas que adotam este tipo de soluções são:

- Redução do “ciclo de vida” dos documentos (devido à melhoria na eficiência das diferentes equipas de trabalho);
- Melhoria da eficiência individual dos funcionários através de:
 - Fácil participação na criação e gestão de documentos em qualquer fase de um projeto;
 - Redução do número de reuniões relativas à documentação;

- Centralização de materiais, que conduz a uma fácil recuperação e gestão de qualquer tipo de documento;
- Inclusão de agentes externos à empresa em determinados projetos (possibilitando o acesso à documentação, de forma pertinente, a outros parceiros, consultores ou clientes);[SB07]

De um modo geral, a eliminação do circuito de papel, a uniformização de metodologias de tratamento de informação, a integração com sistemas de e-mail, fax ou outros e a possibilidade de realizar pesquisas rápidas e eficientes sobre os arquivos digitais representam um forte impacto no funcionamento base das empresas.[HOL]

2.1.2 Geração automática de documentos

O papel da documentação tem ganho cada vez mais importância à medida que as empresas procuram atingir níveis superiores de produtividade com recurso a menos funcionários. Neste tipo de ambiente, o trabalho deve ser cuidadosamente documentado de forma a disponibilizar rapidamente e pertinentemente questões de negócio a clientes e a proporcionar uma melhor gestão interna de informação.[KM94]

A geração automática de documentos leva à criação de uma estrutura coesa que permite impulsionar todas as componentes do fluxo de trabalho das empresas.

Existem inúmeros benefícios associados à geração automática de documentos, nomeadamente:

- Eliminar erros nos documentos;
- Controlo de versões (por exemplo, assegurar que o documento mais atual é distribuído);
- Consolidar múltiplos documentos num único manual;
- Reduzir ou eliminar ciclos de revisão de documentos;
- Melhorar a produtividade e eficiência da empresa;

Para além disto, existem ainda múltiplos benefícios que ajudam indirectamente as empresas que adoptam este tipo de sistemas:

1. Melhoria na precisão da informação:

Para gerar a documentação automaticamente, as bases de dados devem ser precisas e constantemente atualizadas para que os documentos gerados sejam minuciosos. Isto leva a que os utilizadores adaptem o seu comportamento, procedendo de forma a manter sempre toda a informação precisa e atualizada.

2. Aumento da conformidade de dados:

Como resultado da automatização da documentação, as empresas acabam por adquirir melhorias na conformidade de dados, visto que os seus funcionários acedem aos documentos através de uma única plataforma centralizada, podendo apenas gerar documentação a partir da informação presente na mesma.

3. Redução de recursos associados à documentação:

Antes de automatizar a geração de documentos, muitas empresas gastam múltiplos recursos na auditoria e gestão de documentos (não só Recursos Humanos mas também software e tecnologias adicionais). Com a automatização documental, todos estes recursos deixam de ser requeridos, podendo os funcionários ser realocados, sendo também eliminados os custos em tecnologias adicionais.

4. Redução da formação e instrução de novos funcionários:

Em qualquer novo emprego, a componente de trabalho mais complicada e desafiante é a compreensão dos processos internos de uma empresa. Com a geração automática de documentos, as empresas passam a ter processos replicáveis que permitem a rápida integração de qualquer novo funcionário (de forma a que este se possa focar no seu “trabalho”), deixando de existir formação em como criar diferentes tipos de documentação.

5. Proteção de dados confidenciais:

Muitos dos documentos distribuídos internamente ou para o exterior das empresas contêm dados confidenciais - preços, resultados de vendas, dados de clientes, etc. Os processos de geração automática de documentação acabam por limitar a exposição de certas informações (evitando que estas cheguem à pessoa errada), havendo um maior controlo de permissões e entrega de informação.

6. Redução do número de documentos geridos:

Através da geração automática de documentos, passa a existir um controlo elementar das versões da documentação existente e do seu conteúdo, reduzindo o número total de documentos de uma empresa.

7. Elevada consistência de Branding:

Quer seja uma empresa com uma dezena ou milhares de funcionários, um Branding consistente é uma qualidade indispensável. Com a geração automatizada de documentos as empresas têm controlo não só do conteúdo da documentação gerada mas também dos vários templates utilizados, assegurando que a empresa apresenta uma uniformidade no que toca à aparência da sua documentação.

8. Criação de serviços repetíveis:

Uma das principais qualidades de um sistema de geração automática de documentação é o facto de, qualquer funcionário (quer seja um estagiário ou um funcionário que trabalha há anos na empresa) terá acesso à mesma informação e documentação. Como tal, qualquer um poderá criar documentação correta e precisa à medida que é necessário para o fluxo de trabalho da empresa.

2.1.3 Wiki

Uma wiki é um website que pode ser modificado por diferentes utilizadores. Talvez a mais conhecida utilização de uma estrutura de software wiki seja a Wikipedia, uma enciclopédia online que usa o mesmo modelo de desenvolvimento do software open-source. Utilizadores escrevem artigos e colocam-nos na Wikipedia, ficando estes artigos abertos para edição e correção pela comunidade geral de utilizadores da Wikipedia.

No contexto empresarial, tendo em conta o volume de informação processada, este tipo de software é bastante útil devido à flexibilidade e adaptabilidade.[\[kn:h\]](#)

2.1.4 HTML

HTML é a linguagem web centralizada para criar documentos e aplicações para toda a gente utilizar, em qualquer lugar. Hypertext Markup Language (HTML) é o formato usado para construir páginas web e aplicações, e a pedra angular da Plataforma Open Web (uma coleção de tecnologias gratuitas que pertencem à Web, possibilitando o contributo gratuito de todos os utilizadores).[\[Bro07\]](#)[\[kn:m\]](#)

2.1.5 DOCX

Um ficheiro docx é um ficheiro cuja extensão corresponde ao formato Word Microsoft Office Open XML. É uma extensão do Microsoft Word utilizada nas versões a partir do Microsoft Office 2007. Contrariamente às versões antigas do formato word (DOC) o DOCX é um formato pertencente à família dos formatos baseados em XML. As grandes vantagens relativamente às versões mais antigas do formato word são a compatibilidade com diferentes programas e o tamanho menor dos ficheiros.[\[kn:f\]](#)

2.1.6 XML

XML é a sigla para EXtensible Markup Language, sendo um formato baseado em texto que permite a estruturação de documentos eletrónicos. Ao contrário do HTML, o XML é usado para descrever dados (e não apresentá-los) e é uma forma flexível de criar formatos de informação e partilhar dados estruturados via Internet.[\[BD03\]](#)

2.1.7 SharePoint

O SharePoint é uma plataforma colaborativa desenvolvida pela Microsoft que visa suportar as necessidades das empresas através de aplicações web nas componentes: colaborativa, gestão da informação e gestão de processos de negócio. Em todas estas componentes a plataforma inclui tanto a vertente servidor (e respetivas ferramentas de administração), como os serviços para integração de aplicações externas.

Para uma melhor compreensão desta ferramenta e das suas componentes é necessário conhecer alguns conceitos predefinidos desta plataforma, que são parte integrante da sua arquitetura nativa.

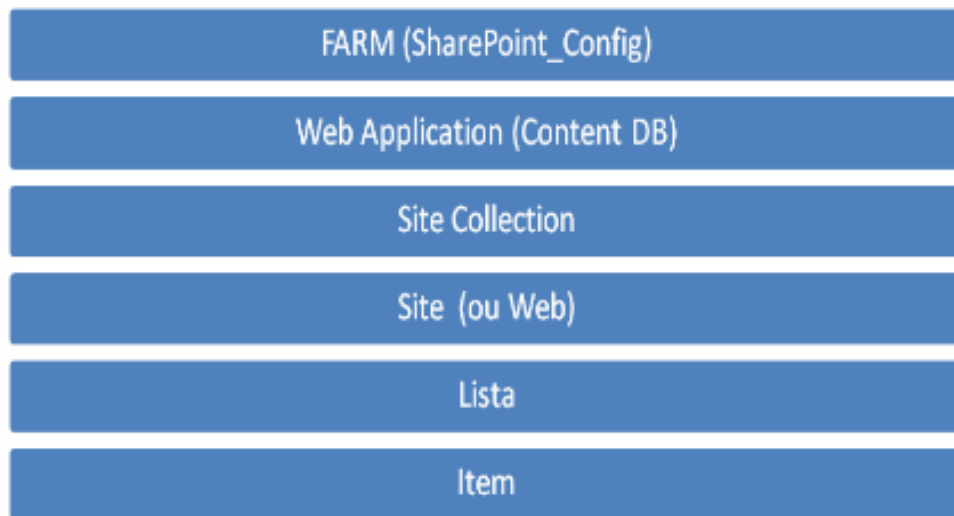


Figura 2.1: Arquitetura Principal do SharePoint

1. Server Farm – elemento de topo na hierarquia do desenho da arquitetura lógica do servidor de SharePoint. É composto por um ou mais servidores instalados e configurados com a ferramenta SharePoint. Num servidor podem estar alojadas diversas Web Applications e/ou Service Applications. Podem existir múltiplas Server Farms.
2. Service Application (Service App) – fornece recursos que podem ser partilhados entre sites dentro de uma Farm ou através de múltiplas Farms. Estão associados às Web Applications através *service application connections*.[\[kn:a\]](#)
3. Web Application (Web App) – Antes de criar uma Site Collection é necessário criar uma Web Application. Cada Web App é composta por uma ou mais Site Collections. Quando criada uma nova Web Application também é gerada uma base de dados(BD) de conteúdo e é definido o método de autenticação usado para se fazer a conexão à BD.[\[kn:d\]](#)
4. Content Database – Fornece capacidade de armazenamento numa única (BD), por default, com o conteúdo da Web Application. Este conteúdo pode ser separado em várias BDs de conteúdo. Este tipo de BD pode conter uma ou mais Site Collections. Uma Site Collection não pode abranger múltiplas BDs. Backups e restauro de sites é a nível das BDs de conteúdo.
5. Site Collection – conjunto de web sites que têm o mesmo “pai” e partilham as mesmas configurações de administração. Cada Site Collection contem um web site de topo e pode conter, também, um ou mais subsites. Para cada site collection existe uma gestão a nível de permissões, navegação e features.
6. Site – constituído por uma ou mais páginas web e outros elementos, como por exemplo, listas, bibliotecas e documentos que estão alojados dentro da Site Collection. Um site pode conter sites “filhos” (embutidos) chamados de subsites. Ambos são essencialmente o mesmo, contudo um subsite vive “por baixo” de outro site.

7. Web Part – é um componente que disponibiliza informação. Pode ser de vários tipos, como por exemplo, branding ou configuração (lista, biblioteca, hiperligação, índice, entre outros), e é colocado ao longo do esquema da página.

[kn:g][SFF13]

2.2 Conversores HTML para DOCX

No âmbito desta tese, realizou-se uma pesquisa aprofundada no que toca ao software disponível no mercado que pudesse facilitar o propósito da ferramenta a desenvolver.

Após realizada uma análise das soluções encontradas, é exposto o motivo pelo qual estas não preenchem todos os requisitos da ferramenta pretendida.

2.2.1 docusuit_

Esta aplicação web permite a criação de documentos personalizados (contratos, relatórios, etc) através de formulários web dinâmicos, de forma rápida e sem conhecimentos de programação.

Os utilizadores do docusuit_ têm apenas de preencher um formulário web automaticamente gerado através de templates predefinidos, de forma a criar os documentos desejados. Esta aplicação permite a criação de documentos Word, PDF, RTF e ODT.

Os custos da aplicação variam de acordo com o número de utilizadores e o número de templates pretendidos, sendo que a opção mais barata é de 1 utilizador e 10 templates (249 euros) e a mais cara de 100 utilizadores e 100 templates (2499 euros) [kn:e]

2.2.2 PHPWord

PHPWord é uma biblioteca open-source, escrita em PHP, que permite criar documentos word. Esta ferramenta tem as seguintes características:

- Inserção e formatação de secções de documentos;
- Inserção e formatação de elementos de Texto;
- Inserção de quebras de Texto;
- Inserção de quebras de Página;
- Inserção de marcas-de-água;
- Inserção de Cabeçalho/Rodapé;
- Inserção e formatação de Tabelas;
- Inserção e formatação de Listas de elementos;
- Inserção e formatação de hiperligações;

Uma das mais-valias desta biblioteca é o facto de ser disponibilizada gratuitamente, para possível integração com outras ferramentas. [\[kn:j\]](#)

2.2.3 Convert Doc

O Convert Doc é um conversor gratuito que permite a conversão de diferentes tipos de ficheiros.

De entre todas as suas componentes destacam-se:

- A possibilidade de utilização da aplicação através da linha de comandos (que corre em background);
- Suporta a conversão rápida de várias extensões de ficheiro - PDF, MS Word (Doc e DocX), HTML, RTF, CSV, TXT, entre outras;
- Possibilidade de utilizar o sistema de conversão original (desenvolvido pela Softinterface) ou o algoritmo de conversão do Microsoft Word;
- Existência de uma agenda para programar futuras conversões automáticas;
- Processamento de ficheiros que inclui concatenação/anexo, remoção de linhas vazias, execução de programas externos, pesquisa e substituição de strings, entre outros;
- Pode ser adaptado a diferentes línguas;

Esta é uma ferramenta prática, rápida e eficaz na conversão de ficheiros. [\[kn:k\]](#)

2.2.4 Conclusões

Após realizado o estudo das características das ferramentas referidas (e experimentação das demais) podemos fundamentar acerca dos motivos pelos quais não são cumpridos todos os requisitos da ferramenta a desenvolver:

- docusuit_:

Embora esta aplicação apresente uma interface bastante intuitiva e pertinente, para além dos elevados custos que esta acarreta, não é permitido que haja uma integração direta no SharePoint (obrigando os funcionários a utilizar duas plataformas distintas), não correspondendo ao pretendido pela Finantech.

- PHPWord:

Esta biblioteca permite a integração com o SharePoint. Contudo, complicaria o processo de conversão relativamente ao método adotado atualmente (que dá uso à biblioteca Open XML para o Office 2013). Concluiu-se que, tirando o facto de o PHPWord obrigar à criação de um Parser mais complexo para obtenção dos mesmos resultados, o Open XML teria funções mais abrangentes para o propósito pretendido.

Trabalho relacionado

- ConvertDoc:

Sendo esta uma ferramenta bastante rápida e eficaz na conversão de documentos, seria uma possível escolha face aos resultados pretendidos. Contudo, relativamente ao processo de conversão, esta iria causar mais complicações quanto à eficiência da ferramenta, dado que só é possível a conversão de documentos integrais (não sendo possível a conversão de conteúdo guardado em strings) e que a sua utilização iria requerer processos externos ao SharePoint a correr em background, executados pela linha de comandos;

Capítulo 3

Wiki2Doc

Neste capítulo é inicialmente referida a motivação associada ao projeto. Posteriormente, é sumariada a arquitetura da ferramenta, incluindo a visão geral do problema, a arquitetura lógica e os casos de uso do utilizador. São também apresentadas as várias tecnologias utilizadas durante o desenvolvimento da ferramenta, assim como o seu enquadramento no projeto. Para além disso faz-se ainda uma análise pormenorizada da implementação do projeto e do desenvolvimento da interface. Por último, são referidas algumas limitações relativas à ferramenta desenvolvida.

3.1 Motivação

No contexto desta dissertação, foi analisado o impacto que este tipo de ferramenta trará para a Finantech, apurando os benefícios adquiridos a partir da sua recorrente utilização. A geração automática de documentos irá permitir à Finantech criar documentos para impressão utilizando os conteúdos que estão guardados na wiki cada vez que for necessário fazer entregas a clientes ou de documentos oficiais. Também irá permitir a eliminação de erros nos documentos (provocados pela falha humana, que se verifica quando os documentos são produzidos manualmente). Para além disto prevê-se que haja uma redução significativa dos recursos associados à documentação (tanto a nível de funcionários responsáveis pela produção e gestão documental como a nível do tempo requerido para este tipo de atividades).

3.2 Requisitos e Restrições da Ferramenta

O levantamento de requisitos é apresentado em dois grupos, o primeiro grupo apresenta os requisitos funcionais, descrevendo explicitamente as funcionalidades e os serviços do sistema, o segundo grupo apresenta os requisitos não-funcionais, definindo propriedades e restrições da ferramenta.

3.2.1 Requisitos Funcionais

Funcionário

- Escolher template que quer utilizar
- Dar um nome ao documento que vai gerar
- Conseguir escolher para cada secção do documento a página de wiki que contém o conteúdo desejado para aquela secção
- Gerar documento

3.2.2 Requisitos Não-Funcionais

Rapidez

A ferramenta deve ser capaz de gerar o documento de uma forma rápida de modo a que o utilizador não gaste muito tempo com esta tarefa.

Consistência

Uma vez que o objetivo da ferramenta é criar um documento pronto para entregar a clientes ou para impressão, é necessário que a transferência de informação seja consistente, mantendo a ordem em que os conteúdos são apresentados, passando imagens e tabelas e ainda mantendo as definições do texto.

Usabilidade

De modo a que o utilizador tenha facilidade em utilizar a Wiki2Doc, esta deve apresentar uma interface intuitiva e simples.

3.3 Arquitetura

De forma a facilmente compreender a estrutura da ferramenta desenvolvida segue-se a explicação pormenorizada da arquitetura presente no sistema.

3.3.1 Visão Geral

No contexto desta dissertação, surge a necessidade de desenvolver uma ferramenta para geração automática de documentação estruturada (manuais, relatórios, etc) utilizando informação contida nas páginas wiki integrada no SharePoint.

Após a inserção de conteúdos na wiki, é dada ao utilizador a hipótese de escolher, para cada secção do documento, o conteúdo da página wiki pretendida, gerando (consoante a seleção de templates predefinidos) a documentação desejada.

No processo de conversão dos conteúdos das páginas wiki para a documentação word, é utilizada uma biblioteca OpenXML especializada para o Office 2013.

Toda a informação do sistema é guardada no SharePoint e em algumas Bases de Dados SQL.

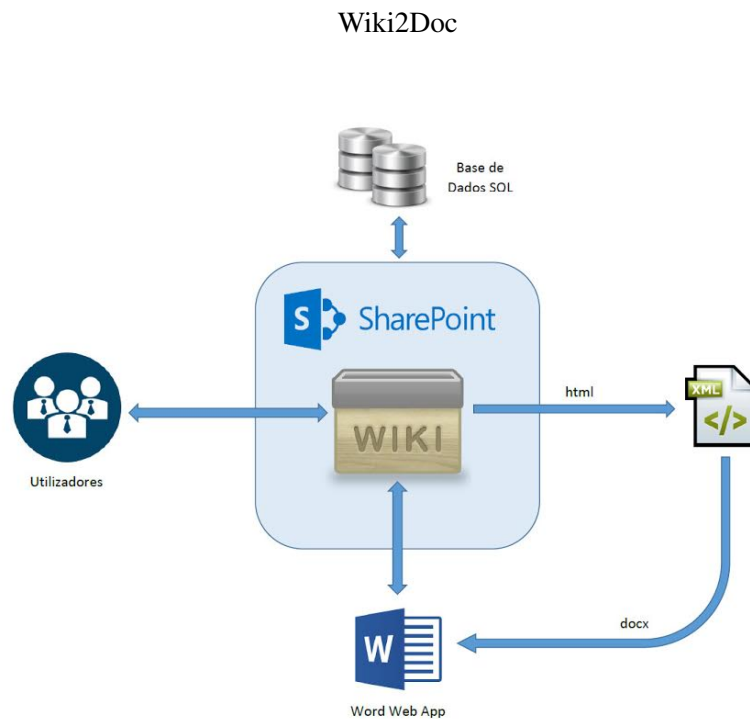


Figura 3.1: Visão Geral

3.3.2 Arquitetura Lógica

A Arquitetura lógica do sistema representa a base lógica do funcionamento da ferramenta desenvolvida. Pretende-se com ela descrever os principais fluxos de informação entre entidades do sistema.

A primeira camada representa a interface do utilizador, que é parte integrante do SharePoint. Esta camada estabelece comunicação com uma segunda camada, a Wiki, que pertence ao sistema de geração automática de documentos. A interface do utilizador recebe ainda informação proveniente de uma terceira camada, a geração de documentação.

Todas as camadas do sistema de geração automática de documentos comunicam com uma outra camada, referente às Bases de Dados.

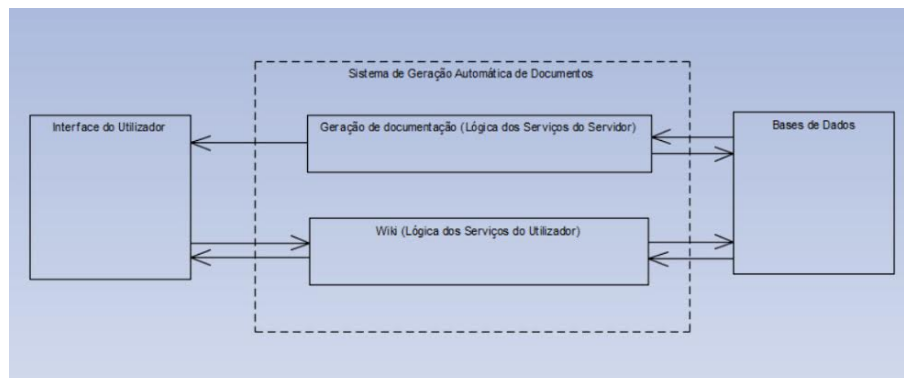


Figura 3.2: Arquitetura Lógica do Sistema

3.3.3 Casos de Uso

Relativamente ao diagrama de casos de uso, parte-se do princípio que todos os utilizadores têm o mesmo nível de acesso e privilégios no que toca à utilização da ferramenta desenvolvida.

Assim sendo, analisando a figura 3.3, pode-se afirmar que o diagrama de casos de uso auxilia a compreensão dos diferentes cenários em que um utilizador tem possibilidade de interagir com o sistema.

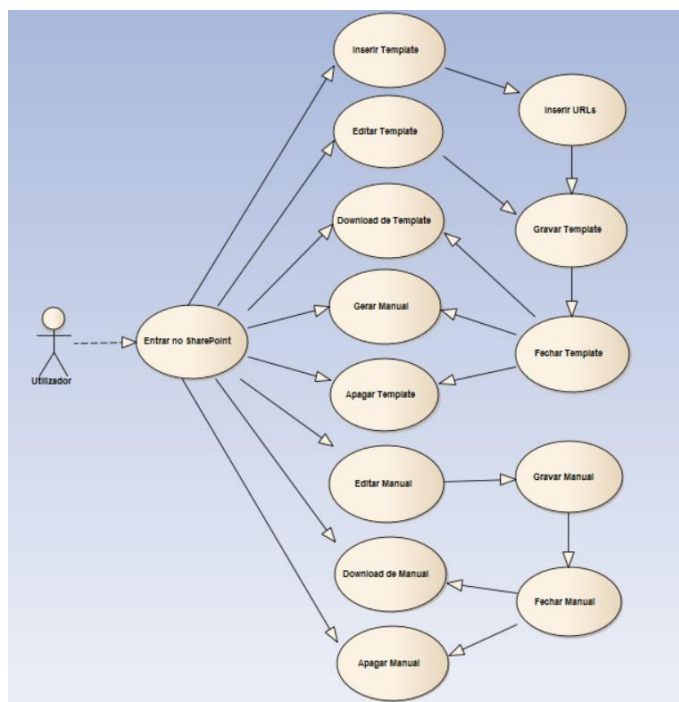


Figura 3.3: Casos de Uso do Utilizador

3.4 Tecnologias usadas

Nos subcapítulos seguintes são introduzidas as ferramentas que, após uma reflexão ponderada, foram selecionadas para a implementação da ferramenta Wiki2Doc.

3.4.1 Open XML SDK 2.5

De modo a conseguir que o conteúdo das páginas wiki fosse convertido para um ficheiro Word utilizou-se a ferramenta OPEN XML SDK 2.5.

O Office Open XML é um formato XML-based desenvolvido pela Microsoft para representar folhas de cálculo, apresentações e documentos de processamento de texto preexistentes que foram codificados em formatos binários definidos pela Microsoft Corporation.

Cada tipo de documento é especificado por uma linguagem markup: WordprocessingML para documentos de processamento de texto, PresentationML para apresentações e SpreadsheetML para folhas de cálculo.[ECM][kn:b]

Estrutura de um Documento WordProcessingML

A estrutura básica de um documento WordProcessingML contém os elementos `<document>` e `<body>`, seguidos de um ou mais parágrafos representados por `<p>`. Um parágrafo contém um ou vários elementos run, representado por `<r>`. Uma região, `<r></r>`, é um bloco de texto que apresenta propriedades comuns, tal como a sua formatação. Dentro de cada região run existem um ou vários elementos `<t>` que são blocos de texto.[6]

A figura seguinte ilustra a linguagem markup de um documento WordprocessingML

```
<w:document xmlns:w="http://schemas.openxmlformats.org/wordprocessingml/2006/main">
  <w:body>
    <w:p>
      <w:r>
        <w:t>Hello, world.</w:t>
      </w:r>
    </w:p>
  </w:body>
</w:document>
```

Figura 3.4: Exemplo da estrutura básica de um documento WordprocessingML

Cada documento WordprocessingML encontra-se organizado em diferentes elementos. Cada elemento é uma região do conteúdo do documento WordprocessingML. Na Figura 3.5 pode-se analisar uma estrutura frequente deste tipo de documentos.[kn:l]

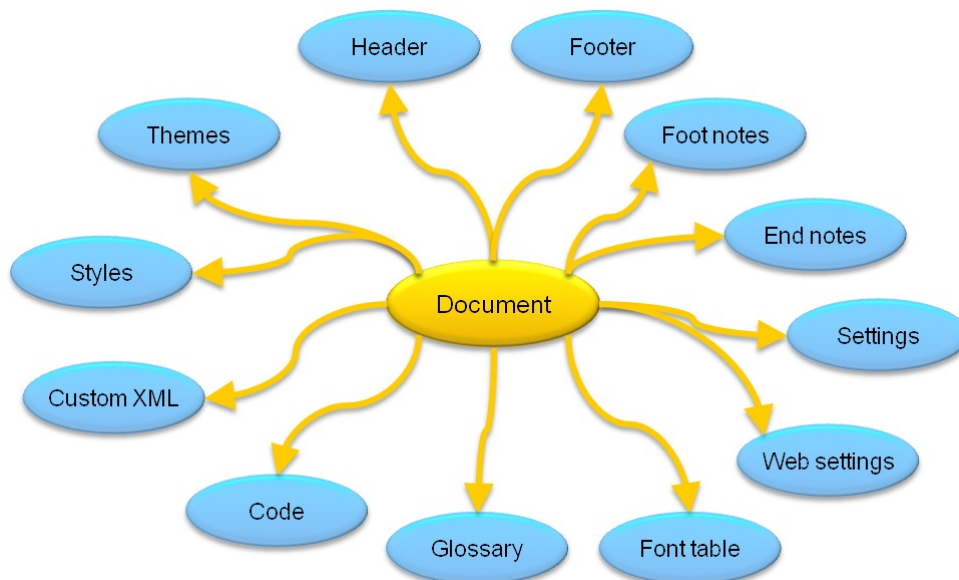


Figura 3.5: Elementos de um documento WordprocessingML

Classe AltChunk

Uma das funções disponíveis na API do Open XML é a classe Althunk (Anchor for Imported External Content). Através dela, consegue-se importar conteúdo de diversos tipos: xml, rtf, texto simples, html, e especificar a localização, dentro do documento, onde vai ser colocado o conteúdo importado para o documento WordProcessingML. Quando é importado conteúdo html é feita a sua conversão para markup WordProcessingML, ou seja, todo o conteúdo que é passado em html é convertido para texto simples, automaticamente.[[kn:c](#)]

3.4.2 HtmlAgilityPack

O “HtmlAgilityPack” é uma biblioteca que permite ao utilizador percorrer ficheiros html. Esta biblioteca transforma conteúdo html numa estrutura em árvore, designada por HTMLDocument. Cada nó da árvore representa uma tag html. O utilizador pode aceder a um HTMLNode e alterar todos os atributos que fazem parte desse nó. É também possível fazer o inverso e converter um HTMLDocument num ficheiro html.[[kn:i](#)]

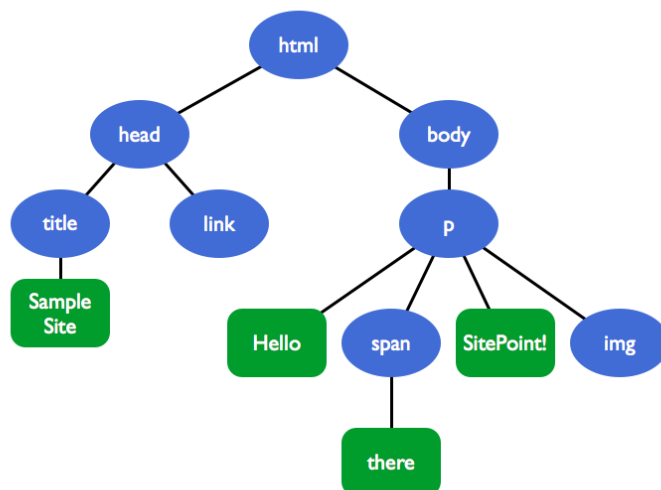


Figura 3.6: Representação HTMLDocument

3.4.3 SharePoint

Como já foi referido no Capítulo 1, a ferramenta usada pela Finantech para gestão documental, sobre a qual foi construída a Wiki2Doc é o SharePoint. Foi também explicado na secção 2.1.7 do Capítulo 2 parte da arquitetura do SharePoint. A figura seguinte representa a estrutura do site da Finantech. Este é constituído por vários subsites (sendo a wiki um deles) e por outros elementos (como por exemplo bibliotecas) e elementos que já pertencem à site collection (nível precedente ao site na arquitetura do SharePoint).

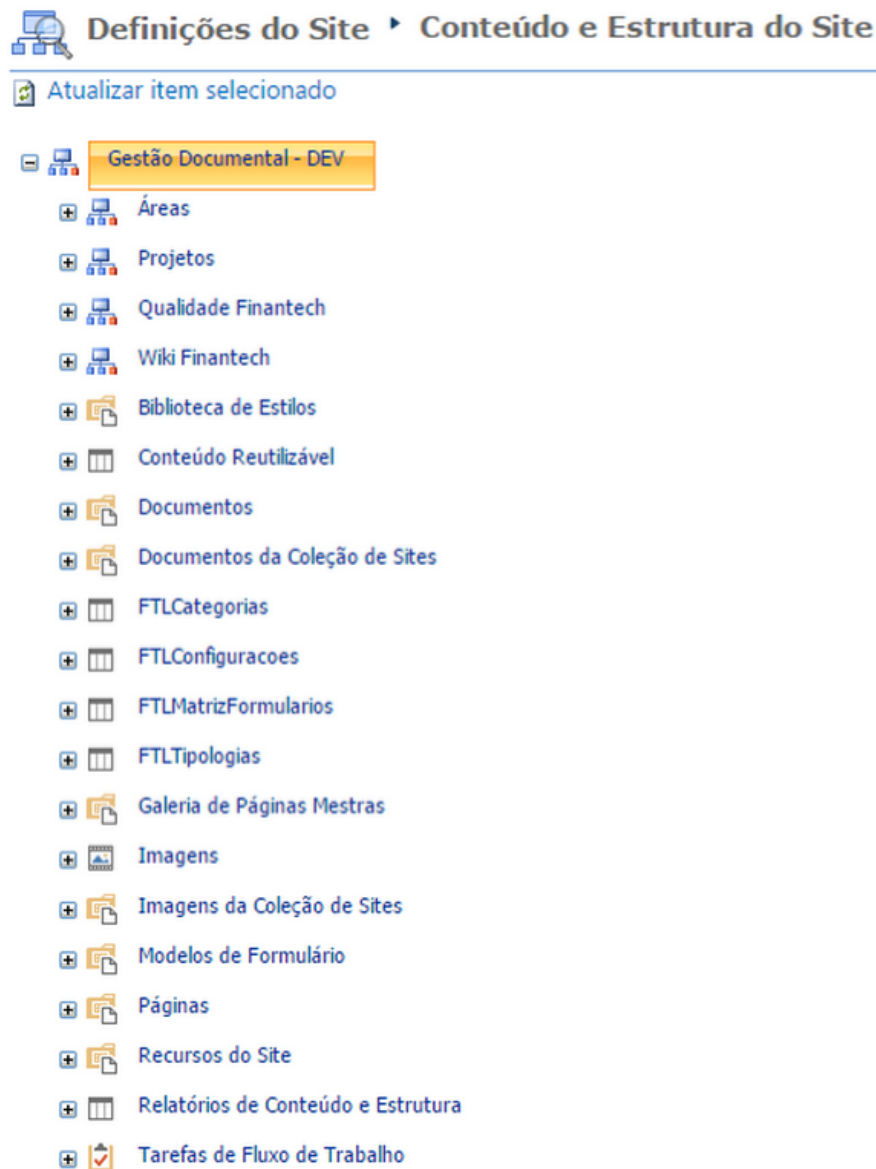


Figura 3.7: Estrutura do Site da Finantech

3.5 Implementação

Após a descrição detalhada da arquitetura e das tecnologias usadas, segue-se uma descrição pormenorizada do desenvolvimento da solução. A solução implementada envolve três fases distintas: A primeira fase corresponde ao desenvolvimento de um conversor que faz a transformação de conteúdo html para Word. A segunda fase corresponde à integração do conversor na ferramenta SharePoint. A terceira fase corresponde ao desenvolvimento da interface, na plataforma SharePoint.

Dado que o projeto se encontra numa fase inicial e que este foi desenvolvido por uma única pessoa, o processo de desenvolvimento utilizado foi o *Modelo em Cascata (Waterfall Model)* uma

vez que a sua característica sequencial se adequava mais ao projeto nas condições existentes.

3.5.1 Conversor de Html para Word

De modo a fazer a conversão do conteúdo html para docx foi elaborada uma função chamada *AltChunkMethod()* que recebe um *WordprocessingDocument*, uma string (que contém o html extraído da página wiki) e o id do altchunk (que é também uma string). Esta função cria um elemento *XElement altChunk* que é colocado numa zona específica do documento, sendo o conteúdo da página wiki inserido no ficheiro a partir dessa posição.

```
1 void AltChunkMethod(WordprocessingDocument outDoc, string htmlbuilder, string
    chunkid)
```

```
<w:body>
<w:altChunk r:id="altChunk1" />
<w:p/>
<w:sectPr>
...
</w:sectPr>
</w:body>
```

Figura 3.8: Representação do elemento XElement do tipo Altchunk

O elemento Altchunk só consegue transferir informação em "blocos", ou seja, o conteúdo que lhe é passado será integralmente colocado no novo ficheiro. Esta particularidade é a razão pela qual o seu tempo de execução é muito reduzido, o que permite que o conversor seja bastante eficaz.

3.5.2 Integração do conversor no SharePoint

Depois de implementado o conversor, foi feita a sua integração no SharePoint.

De forma a obter o conteúdo de uma página wiki foi desenvolvida uma função chamada *GetWikiPageHTML()* que retorna uma string com todo o conteúdo html:

```
1 string GetWikiPageHTML(string site_url, string listName, string pageName)
```

Uma vez extraído o html, foram encontrados alguns obstáculos na passagem de imagens dado que o seu atributo "src" não incluía todo o url de referência. Surgiu então a necessidade de usar a biblioteca *HTMLAgilityPack* para se conseguir encontrar todas as imagens presentes na string e modificar o seu url. Para este efeito foi desenvolvida uma função chamada *FindHTMLImages()*

que recebe uma string com o html da página wiki, transformando-o num HTMLDocument (onde são modificados todos os url's das imagens), e voltando a transformá-lo numa string:

```
1 string FindHTMLImages(string html)
```

Depois de integrado o conversor no SharePoint foi necessário começar a pensar na forma como seriam colocados os conteúdos das várias páginas wiki nas diferentes posições do documento a gerar.

Cada documento é criado a partir de um template predefinido, que já está inserido no Share-Point (Anexo A - Figuras A1, A2, A3, A4, e A10, A11, A12, A13) e a wiki da empresa encontra-se organizada por secções e subsecções (Figura 3.9, Figura 3.10 e Figura 3.11), de modo a que cada página corresponda a uma secção de um template. Como tal, chegou-se à conclusão de que a melhor forma de associar os conteúdos aos templates seria colocando o url da página desejada diretamente abaixo do título da secção do template.

Assim, como se pode observar na figura 3.12., cada secção do template a utilizar terá um url (placeholder) ao qual é associada a informação da página wiki que lhe corresponde.



Figura 3.9: Índice da wiki integrada no SharePoint (1º Nível)



Figura 3.10: Índice da wiki integrada no SharePoint (2º Nivel - Ferramenta SOAPP)

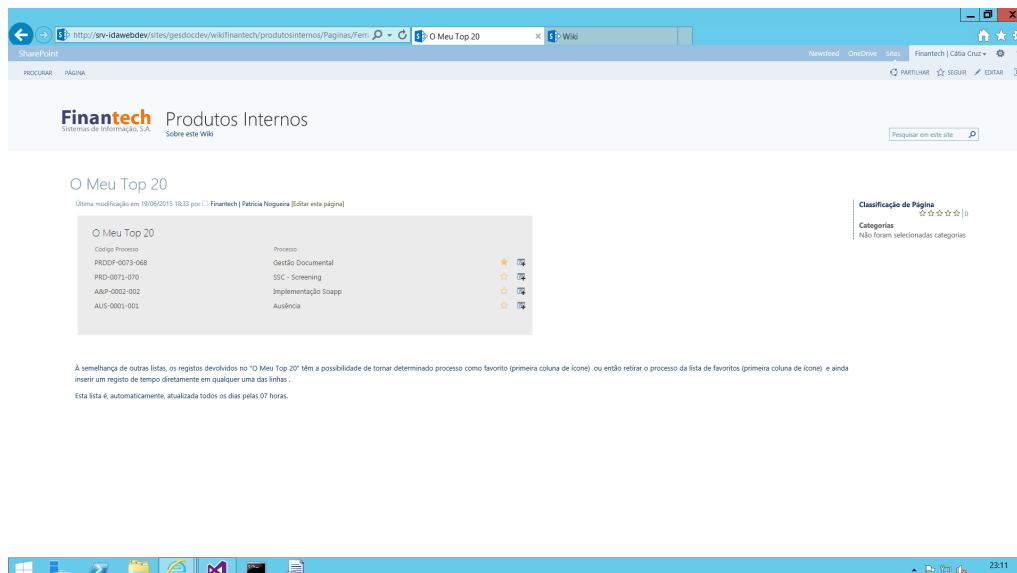


Figura 3.11: Exemplo de página wiki integrada no SharePoint (“O Meu Top 20”)

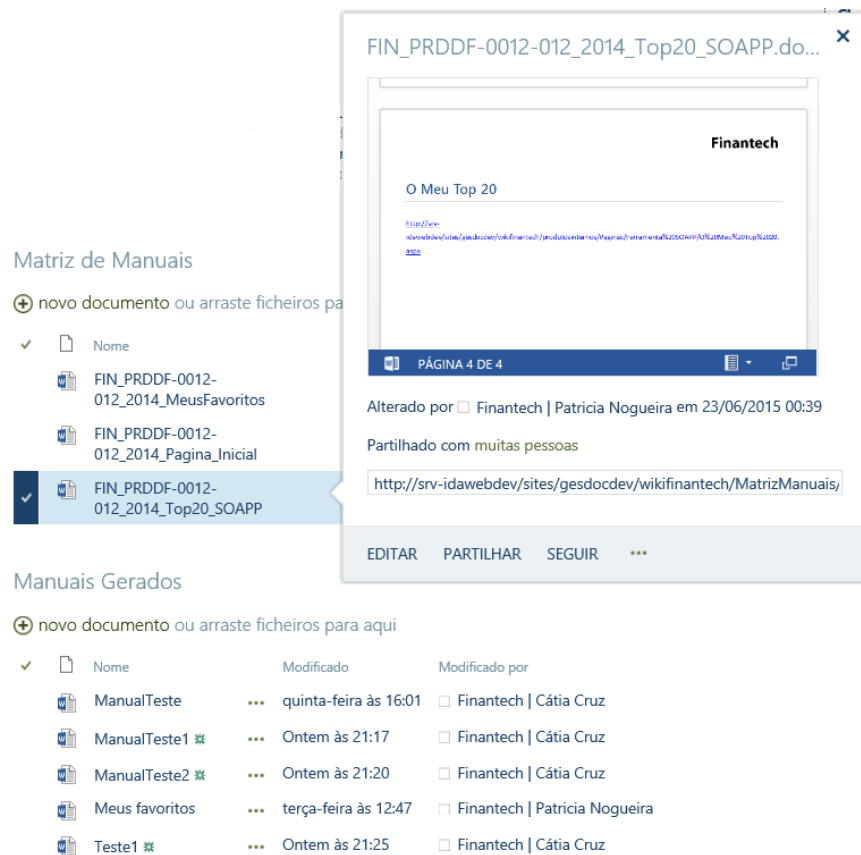


Figura 3.12: Template com secção correspondente à página “O Meu Top 20”

Tendo em conta que o url das páginas wiki é utilizado como placeholder para os templates, é possível obter os três parâmetros a inserir na função *GetWikiPageHTML()*, de forma a extrair o conteúdo da respetiva página.

De modo a extrair o conteúdo de todas as páginas wiki correspondentes às várias secções de um template, foi implementada a função cíclica, com o número de ciclos igual ao número de url's presentes no template. *btnGenerateManual_Click()*.

```
1 void btnGenerateManual_Click(object sender, System.Web.UI.ImageClickEventArgs e)
```

Esta função percorre todo o template e, cada vez que encontra um url, chama a função *GetWikiPageHTML()* para extrair o conteúdo da página, que é guardado numa string, que por sua vez é passado como parâmetro na função *AltChunkMethod()* de forma a escrever o conteúdo no documento a gerar. No final os documentos (Anexo A - Figuras A5, A6, A7, A8, A9 e A14, A15, A16, A17) gerados são guardados no SharePoint. A figura seguinte exemplica todo o funcionamento da ferramenta.

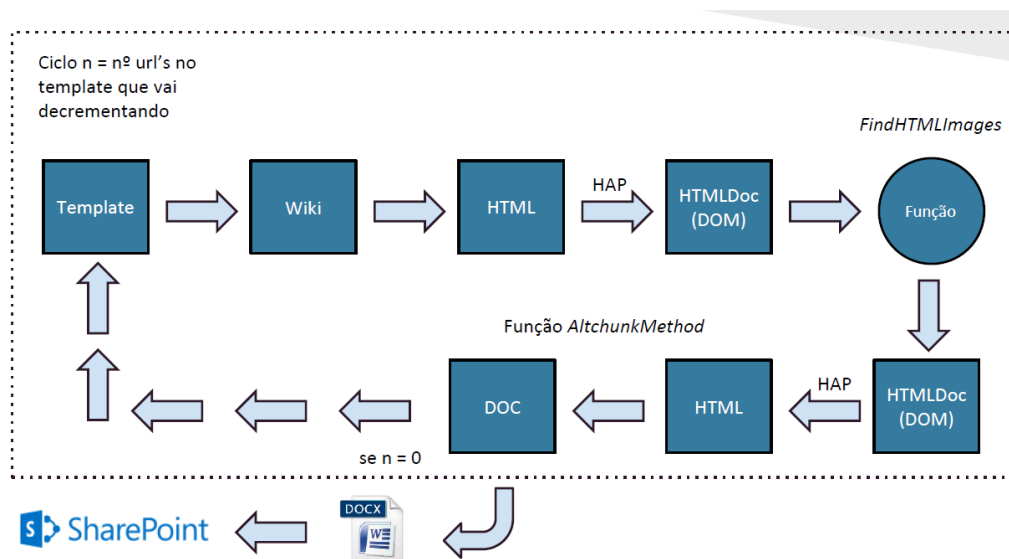


Figura 3.13: Fluxo de passagem da informação

3.6 Desenvolvimento da Interface

Os métodos anteriores foram desenvolvidos sob uma interface, pelo que necessitam da entrada de dados por parte do utilizador. Seguidamente, é apresentado um fluxo das várias etapas que o utilizador deve efetuar, baseado no diagrama de casos de uso apresentado no Capítulo 3.

Inicialmente, o utilizador deve seleccionar a opção “Wiki Finantech” do menu principal, representado pela figura 3.14.

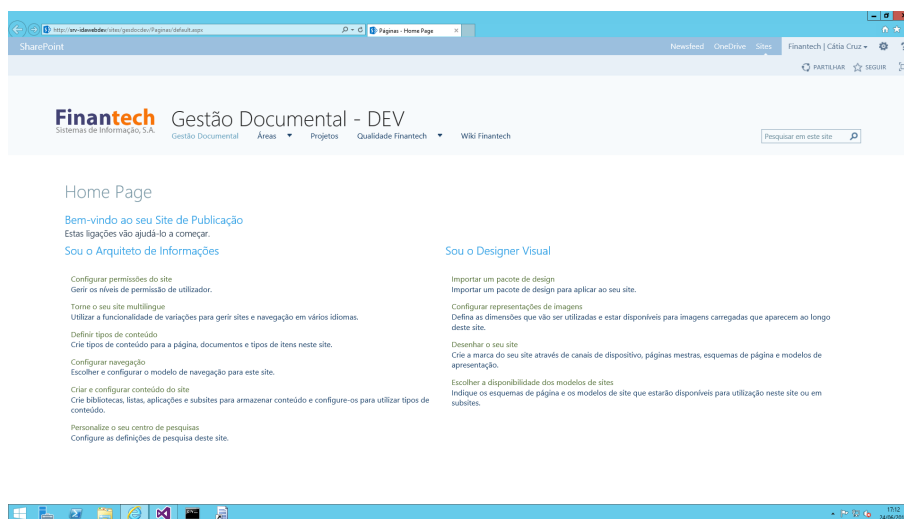


Figura 3.14: Página inicial do SharePoint

Após a seleção desta opção no menu, o utilizador consegue visualizar a página da wiki (Figura 3.15).

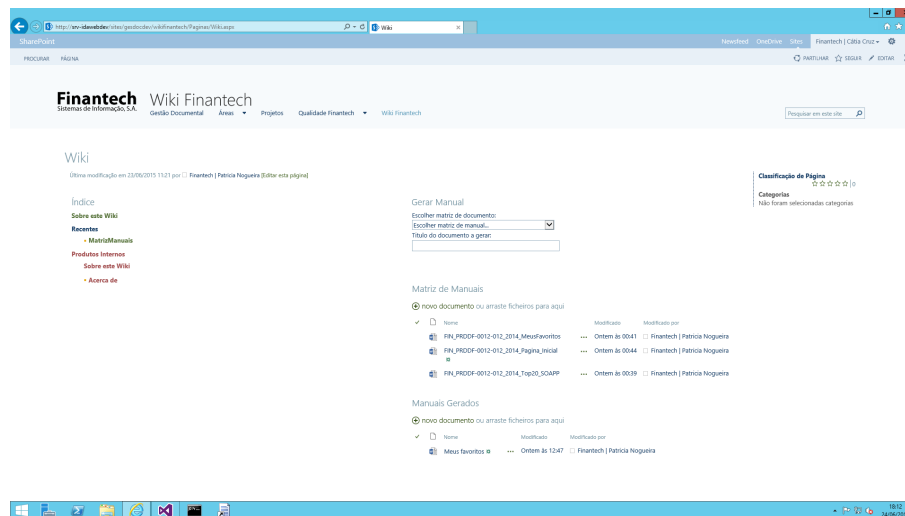


Figura 3.15: Página Inicial da Wiki integrada no SharePoint

De seguida, o utilizador deve aceder ao menu “Matriz de Manuais”(Figura 3.16), seleccionar o template que deseja e escolher a opção de edição. Durante a edição, deve-se introduzir por baixo do título de cada secção do template o url da página wiki com o conteúdo que se pretende colocar nessa parte do documento.

Matriz de Manuais

+ novo documento ou arraste ficheiros para aqui

| ✓ | Nome | Modificado | Modificado por |
|---|--|------------|---|
| | FIN_PRDDF-0012-012_2014_MeusFavoritos | ... | Ontem às 00:41 Finantech Patricia Nogueira |
| | FIN_PRDDF-0012-012_2014_Pagina_Inicial | ... | Ontem às 00:44 Finantech Patricia Nogueira |
| | FIN_PRDDF-0012-012_2014_Top20_SOAPP | ... | Ontem às 00:39 Finantech Patricia Nogueira |

Figura 3.16: Menu “Matriz de Manuais”

Depois de introduzir os url’s, o utilizador deve escolher, no menu “Gerar Manual”(Figura 3.17), o template (matriz de documento) que deseja utilizar e dar um título ao documento que vai ser gerado.

Após serem preenchidos, obrigatoriamente, ambos os campos, o utilizador deve clicar no botão, sendo gerado um novo documento que será colocado no menu “Manuais Gerados”(Figura 3.18).

Gerar Manual

Escolher matriz de documento:

FIN_PRDDF-0012-012_2014_MeusFavoritos.docx

Título do documento a gerar:

Manual1

Clicar aqui para gerar manual

Figura 3.17: Menu “Gerar Manual”

Manuais Gerados

+ novo documento ou arraste ficheiros para aqui

| ✓ | Nome | Modificado | Modificado por |
|----------------|------|----------------|-------------------------------|
| Meus favoritos | ... | Ontem às 12:47 | Finantech Patricia Nogueira |

Figura 3.18: Menu “Manuais Gerados”

3.7 Limitações da Ferramenta

Uma vez que todo o projeto de gestão documental é muito recente, a ferramenta de geração automática de documentação possui ainda algumas limitações, tanto a nível de interface do utilizador como a nível operacional. Apresentam-se abaixo algumas das limitações identificadas.

3.7.1 Interface na edição de Templates

Nesta fase do projeto, um utilizador, para gerar um documento tem de colocar (manualmente), no template, os url's das páginas wiki desejadas.

Falta ainda desenvolver uma interface intuitiva e user-friendly, que permita criar correspondência entre as secções dos templates e as páginas wiki existentes.

3.7.2 Controlo de Versões

Provisoriamente, o projeto não tem qualquer tipo de controlo de versões no que toca aos templates e aos documentos gerados. Quando um utilizador edita um template ou documento, este é guardado no sistema, perdendo todas as versões anteriores.

3.7.3 Estrutura das secções dos Templates

Tendo em conta a atual estrutura da wiki, o sistema de secções dos templates é bastante limitado, dado que cada secção corresponde obrigatoriamente a uma página completa da wiki. A ferramenta é adequada ao atual contexto organizacional da wiki da empresa mas não está otimizada para contextos mais generalizados (como por exemplo, a possibilidade de selecionar pequenos excertos de diferentes páginas da wiki para a mesma secção de um template).

3.7.4 Dificuldades Sentidas

Ao longo do desenvolvimento do projeto, foram superadas diversas dificuldades encontradas em diferentes fases. Tendo em conta a elevada complexidade da arquitetura do SharePoint, foi trabalhoso desenvolver o código que extrai o conteúdo html das páginas da wiki. Durante a integração do conversor no SharePoint, surgiram bastantes complicações no que toca à conversão de imagens. No processo de resolução deste problema foi ponderada a implementação de um parser para modificar o url das imagens (que vinha incompleto) mas optou-se pela utilização do HTMLAgilityPack que já permitia aceder a diferentes atributos das imagens (nomeadamente o seu url). Na fase de testes, uma vez que os servidores estavam sobrecarregados, tornou-se complicado testar a ferramenta pois não era possível aceder aos conteúdos guardados no SharePoint.

3.8 Conclusões Retiradas

Finalizada a explicação de todo o processo de análise da Wiki2Doc, pode-se concluir que, apesar da existência de algumas restrições na sua usabilidade, a ferramenta corresponde consideravelmente às necessidades requeridas pela Finantech. Haverá ainda bastantes reformulações e melhorias (características de qualquer projecto nesta fase) contudo, neste momento, a Finantech já consegue produzir bastantes manuais que se enquadrem na estrutura da ferramenta implementada. Relativamente ao processo de desenvolvimento adotado (o modelo em cascata), caso existisse uma equipa de trabalho com um maior número de elementos, teria sido possível a utilização de um processo mais sofisticado e produtivo como por exemplo o agile ou o scrum. Futuramente, é aconselhada a adoção de um destes processos de desenvolvimento.

Capítulo 4

Conclusões e Trabalho Futuro

O desenvolvimento da ferramenta Wiki2Doc foi vantajoso para compreender o contexto atual da gestão documental da Finantech. Desta forma é possível retirar algumas conclusões face ao estado corrente do projeto e possível trabalho a realizar no futuro.

4.1 Conclusões

Os sistemas de gestão documental assumem um papel essencial no bom funcionamento de quaisquer empresas. Como tal, a otimização destes sistemas conduz a melhorias significativas em todos os fluxos de trabalho, poupando tempo e recursos. A geração automática de documentos é um dos componentes que mais exponencia os benefícios experienciados pelas empresas que se dedicam à melhoria dos seus processos de gestão documental, produzindo inúmeras vantagens, não só ao nível interno mas também na forma como estas são vistas pelo exterior.

A Finantech é uma empresa que iniciou recentemente um projeto de desenvolvimento de uma plataforma de documentação que consiste num centro de documentos e numa wiki. Embora este projeto esteja numa fase bastante prematura, as perspetivas futuras correspondem a progressos relevantes na organização interna da empresa, que, por sua vez, irão potenciar a posição no mercado em que esta se encontra inserida.

A ferramenta Wiki2Doc possibilita a geração automática de documentação estruturada, através da transferência dos conteúdos presentes na wiki da Finantech para os documentos a gerar, demonstrando-se uma ferramenta bastante eficiente graças às tecnologias utilizadas. Desta forma, podemos concluir, que a escolha do OpenXML para este projeto em específico permitiu que os conteúdos das páginas da wiki fossem transmitidos de forma rápida (devido à passagem da informação em blocos - strings) tornando o conversor bastante vantajoso no que toca à estrutura organizacional de informação. Para além da ágil e eficiente transmissão de dados, o conversor tem a capacidade de preservar todos os conteúdos, convertendo corretamente os vários formatos de informação (texto, imagens, tabelas, etc) de html para documento Word (docx).

No contexto do mercado atual, a ferramenta que foi desenvolvida encontra-se ainda em fase de prototipagem, demonstrando algumas limitações típicas de um produto que se encontra em fase

de maturação. Uma das principais limitações é a sua interface pouco desenvolvida que obriga os utilizadores a terem de aceder ao template e colocar manualmente, nos locais respetivos, os links para das páginas que contém a informação que será colocada naquela secção do documento. Para além deste problema, a Wiki2Doc também não permite gravar diversas versões de um template, ou seja, se o utilizador modificar o template e apagar alguma secção que não seja necessária para o documento final que quer gerar, não consegue recuperar o template na sua forma original. No entanto, face a outras aplicações já disponíveis no mercado, a Wiki2Doc apresenta fortes mais-valias, sendo completamente integrada no SharePoint e, do ponto de vista da Finantech, cem por cento projetada para satisfazer as necessidades da empresa.

4.2 Trabalho Futuro

Como continuação do trabalho desenvolvido nesta dissertação, prevê-se que a Finantech irá focar-se fortemente na interface da ferramenta, visto que há ainda bastantes melhorias a efetuar. Após um estudo cuidado e atencioso, a interface poderá ser projetada para ser mais autónoma, intuitiva e user-friendly, evitando, por exemplo, a necessidade de os utilizadores inserirem manualmente informação nos templates (passando a fazê-lo mediante a seleção de opções predefinidas, que variassem consoante os conteúdos disponibilizados na wiki).

Uma das funcionalidades a implementar futuramente foca-se ainda na possibilidade de gravar várias versões do mesmo template (para fornecer a diferentes clientes, por exemplo) ou mesmo reaver uma versão antiga de um template ou de um documento previamente gerado. Por último, sugere-se ainda a validação de uma renovação no que toca à estrutura organizacional da informação (tanto dos templates como da wiki), possibilitando a existência de mais do que um tipo de secções, tornando a documentação mais flexível e adaptável a novos requisitos.

Referências

- [BD03] Brain Benz e John R. Durant. *XML Programming Bible*. Wiley Publishing Inc., 2003.
- [Bro07] David R. Brookst. *An Introduction to HTML and JavaScript: for Scientists and Engineers*. Springer, 2007.
- [ECM] ECMA. Visão geral do office open xml. Technical report, ECMA.
- [HOL] HOLOS. Holdoc - gestao documental.
- [KM94] James Shaw Kathleen McKeown, Karen Kukich. Practical issues in automatic documentation generation. 1994.
- [kn:a] About service applications and services in sharepoint 2013. Disponível em <https://technet.microsoft.com/en-us/library/cc261875.aspx>.
- [kn:b] About the open xml sdk 2.5 for office. Disponível em <https://msdn.microsoft.com/EN-US/library/office/bb456487.aspx>.
- [kn:c] Altchunk class. Disponível em [https://msdn.microsoft.com/en-us/library/documentformat.openxml.wordprocessing.altchunk\(v=office.14\).aspx](https://msdn.microsoft.com/en-us/library/documentformat.openxml.wordprocessing.altchunk(v=office.14).aspx).
- [kn:d] Create a web application in sharepoint 2013. Disponível em <https://technet.microsoft.com/en-us/library/cc261875.aspx>.
- [kn:e] docusuit. Disponível em <http://www.docusuit.com/>.
- [kn:f] ehow. Disponível em http://www.ehow.com/facts_5733176_file-extension-docx-definition.html.
- [kn:g] Elements to structure sharepoint. Technical report.
- [kn:h] Encyclopedia britannica. Disponível em <http://www.britannica.com/topic/wiki>.
- [kn:i] Html agility pack. Disponível em <https://htmlagilitypack.codeplex.com/>.
- [kn:j] Phpword. Disponível em <https://phpword.codeplex.com/>.
- [kn:k] Softinterface - convert doc. Disponível em <http://www.softinterface.com/Convert-Doc/Convert-Doc.htm>.
- [kn:l] Structure of a wordprocessingml document (open xml sdk). Disponível em <https://msdn.microsoft.com/en-us/library/office/gg278308.aspx>.

REFERÊNCIAS

- [kn:m] W3c - html, the web's core language. Disponível em <http://www.w3.org/html/>.
- [SB07] Tony Smith e Seth Bates. *SharePoint 2007 User's Guide: Learning Microsoft's Collaboration and Productivity Platform*. Apress, 2007.
- [SFF13] Chris Johnson Steve Fox e Donavan Follette. *Beginning SharePoint 2013 Development*. John Wiley and Sons Inc., 2013.
- [Teca] Drawloop Technologies. Drawloop technologies. Disponível em <http://www.drawloop.com/blog/2015/01/21/eight-indirect-benefits-of-an-automated-document-generation-process-part-one/>.
- [Tecb] Drawloop Technologies. Drawloop technologies. Disponível em <http://www.drawloop.com/blog/2015/01/21/eight-indirect-benefits-of-an-automated-document-generation-process-part-two/>.

Anexo A

Templates e Manuais Gerados



Manual de Utilizador – Página Inicial SOAPP

Finantech – Sistemas de Informação SA
Av. Fernão de Magalhães, 1862, 15º andar
4350-158 Porto Portugal
Telefone: +351 22 507 20 30
Fax: +351 22 507 20 59
Email: business@finantech.pt

Figura A.1: Template de Exemplo 1 (Pág.1)

Revisões ao Documento

| Data | Razão/Comentário | Versão | Versão | Responsável |
|------------|-------------------------------|--------|--------------|-------------|
| 18-16-2015 | Manual – Página Inicial SOAPP | 1.2 | Renata Silva | Pedro Silva |

Figura A.2: Template de Exemplo 1 (Pág.2)

Finantech

Índice

| | |
|----------------------------|---|
| Índice..... | 3 |
| Página Inicial SOAPP | 4 |

Manual de Utilizador

3 de 4

Figura A.3: Template de Exemplo 1 (Pág.3)

Página Inicial SOAPP

(1)

<http://srv->

idawebdev/sites/gesdocdev/wikifinantech/produtosinternos/Paginas/Ferramenta%20SOAPP/P%C3%A1gina%20Inicial%20-%20SOAPP.aspx

Figura A.4: Template de Exemplo 1 (Pág.4)



Manual de Utilizador – Página Inicial SOAPP

Finantech – Sistemas de Informação SA
Av. Fernão de Magalhães, 1862, 15º andar
4350-158 Porto Portugal
Telefone: +351 22 507 20 30
Fax: +351 22 507 20 59
Email: business@finantech.pt

Figura A.5: Manual de Exemplo 1 (Pág.1)

Finantech

Revisões ao Documento

| Data | Razão/Comentário | Versão | Versão | Responsável |
|------------|-------------------------------|--------|--------------|-------------|
| 18-16-2015 | Manual – Página Inicial SOAPP | 1.2 | Renata Silva | Pedro Silva |

Figura A.6: Manual de Exemplo 1 (Pág.2)

Índice

Finantech

| | |
|----------------------------|---|
| Índice..... | 3 |
| Página Inicial SOAPP | 4 |

Figura A.7: Manual de Exemplo 1 (Pág.3)

Página Inicial SOAPP

Na intranet, a partir da hiperligação disponível na secção "Serviços Internos", o utilizador, com login efetuado, terá acesso à página onde se encontra uma nova ferramenta cujo objetivo é o de facilitar a criação de registos no SOAPP.

| Serviços Internos |
|---|
| ▶ Site de Apoio ao Cliente (SAC) |
| ▶ SOAPP |
| ▶ SOAPP (2014) - Somente Para Consulta |
| ▶ Simax |
| ▶ Sifox Launcher |
| ▶ Queries do Serviço de Monitorização |
| ▶ Registo de Tempos no Soap |
| ▶ Relatório - Tempos de Trabalho |

Nesta ferramenta, o utilizador terá acesso às seguintes funcionalidades:

- O Meu Top 20;
- Os Meus Favoritos;
- Os Meus Registos no SOAPP;
- Pesquisa de Processos;
- Registos de Tempos a enviar para o SOAPP.

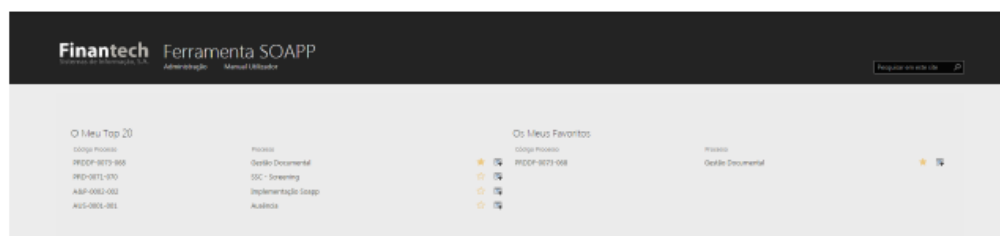


Figura A.8: Manual de Exemplo 1 (Pág.4)

Templates e Manuais Gerados

Finantech

Pesquisar Processo

Cliente:
 Registro:
 Tipo Processo:
 Nome processo:
 Código Processo:

Filtrar na tabela:

| Código Processo | Processo | Cliente | Acronímico | Registro |
|----------------------------------|----------|---------|------------|----------|
| Não foram encontrados resultados | | | | |

Seu registro para mostrar: Anterior Seguinte

Os Meus Registros no SOAPP

• junho de 2015

Expandir Tudo | | |

| DOMINGO | SEGUNDA-FEIRA | TERÇA-FEIRA | QUARTA-FEIRA | QUINTA-FEIRA | SEXTA-FEIRA | SÁBADO |
|---------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| 31 | 30 | 29 | 28 | 27 | 26 | 25 |
| | REG-0012 REG-0013-008 | REG-0014 REG-0015-008 | REG-0016 REG-0017-008 | REG-0018 REG-0019-008 | REG-0020 REG-0021-008 | REG-0022 REG-0023-008 |
| 7 | 6 | 5 | 4 | 3 | 2 | 1 |
| | REG-0024 REG-0025-008 | REG-0026 REG-0027-008 | REG-0028 REG-0029-008 | REG-0030 REG-0031-008 | REG-0032 REG-0033-008 | REG-0034 REG-0035-008 |
| 14 | 13 | 12 | 11 | 10 | 9 | 8 |
| | REG-0036 REG-0037-008 | REG-0038 REG-0039-008 | REG-0040 REG-0041-008 | REG-0042 REG-0043-008 | REG-0044 REG-0045-008 | REG-0046 REG-0047-008 |
| 21 | 20 | 19 | 18 | 17 | 16 | 15 |
| | REG-0048 REG-0049-008 | REG-0050 REG-0051-008 | REG-0052 REG-0053-008 | REG-0054 REG-0055-008 | REG-0056 REG-0057-008 | REG-0058 REG-0059-008 |
| 28 | 27 | 26 | 25 | 24 | 23 | 22 |
| | REG-0060 REG-0061-008 | REG-0062 REG-0063-008 | REG-0064 REG-0065-008 | REG-0066 REG-0067-008 | REG-0068 REG-0069-008 | REG-0070 REG-0071-008 |

Registros de Tempos a enviar para o SOAPP

| Cliente | Processo | Atividade | DATA/HORA INÍCIO | DATA/HORA FIM | Assunto | Tempo Dedicado | | | |
|---|----------------------|------------------------|------------------|------------------|--------------------------|----------------|--|--|--|
| Data: 16/04/2015 | | | | | | | | | |
| Corpo Banco de Investimentos, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 16/04/2015 09:00 | 16/04/2015 09:00 | Atribuição do Projeto... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | ASP-Atividade | 16/04/2015 09:00 | 16/04/2015 10:00 | Atividade no Trabalho | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Interação Mensagem | 16/04/2015 09:00 | 16/04/2015 09:00 | Verificação do Vêni... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Atividade | 16/04/2015 10:00 | 16/04/2015 10:00 | Procedimento para o... | 01h:00m | | | |
| Data: 17/04/2015 | | | | | | | | | |
| Corpo Banco de Investimentos, S.A. | (H4P-0041-01) CAD... | H4P-Interação Mensagem | 17/04/2015 09:00 | 17/04/2015 09:00 | Verificação do Vêni... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 17/04/2015 09:00 | 17/04/2015 10:00 | Atribuição do Projeto... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Interação Mensagem | 17/04/2015 10:00 | 17/04/2015 10:00 | Verificação do Vêni... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 17/04/2015 10:00 | 17/04/2015 10:00 | Procedimento para o... | 01h:00m | | | |
| Corpo Banco de Investimentos, S.A. | (H4P-0041-01) CAD... | H4P-Atividade | 17/04/2015 10:00 | 17/04/2015 10:00 | Procedimento para o... | 01h:00m | | | |
| Data: 18/04/2015 | | | | | | | | | |
| Corpo Banco de Investimentos, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 18/04/2015 09:00 | 18/04/2015 09:00 | Verificação do Vêni... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 18/04/2015 09:00 | 18/04/2015 10:00 | Atribuição do Projeto... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Interação Mensagem | 18/04/2015 10:00 | 18/04/2015 10:00 | Verificação do Vêni... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 18/04/2015 10:00 | 18/04/2015 10:00 | Procedimento para o... | 01h:00m | | | |
| Corpo Banco de Investimentos, S.A. | (H4P-0041-01) CAD... | H4P-Atividade | 18/04/2015 10:00 | 18/04/2015 10:00 | Procedimento para o... | 01h:00m | | | |
| Data: 19/04/2015 | | | | | | | | | |
| Corpo Banco de Investimentos, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 19/04/2015 09:00 | 19/04/2015 09:00 | Verificação do Vêni... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 19/04/2015 09:00 | 19/04/2015 10:00 | Atribuição do Projeto... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Interação Mensagem | 19/04/2015 10:00 | 19/04/2015 10:00 | Verificação do Vêni... | 01h:00m | | | |
| Finantech - Sistema de Informação, S.A. | (H4P-0041-01) CAD... | H4P-Desenvolvimento | 19/04/2015 10:00 | 19/04/2015 10:00 | Procedimento para o... | 01h:00m | | | |
| Corpo Banco de Investimentos, S.A. | (H4P-0041-01) CAD... | H4P-Atividade | 19/04/2015 10:00 | 19/04/2015 10:00 | Procedimento para o... | 01h:00m | | | |

Figura A.9: Manual de Exemplo 1 (Pág.5)



Manual de Utilizador – Top 20

Finantech – Sistemas de Informação SA
Av. Fernão de Magalhães, 1862, 15º andar
4350-158 Porto Portugal
Telefone: +351 22 507 20 30
Fax: +351 22 507 20 59
Email: business@finantech.pt

Figura A.10: Template de Exemplo 2 (Pág.1)

Finantech

Revisões ao Documento

| Data | Razão/Comentário | Versão | Versão | Responsável |
|------------|------------------|--------|--------------|-------------|
| 18-16-2015 | Top 20 - Manual | 1.0 | Renata Silva | Pedro Silva |

Figura A.11: Template de Exemplo 2 (Pág.2)

Finantech

Índice

Índice.....3

O Meu Top 204

Manual de Utilizador

3 de 4

Figura A.12: Template de Exemplo 2 (Pág.3)

Finantech

O Meu Top 20

(1)

<http://srv->

idawebdev/sites/gesdocdev/wikifinantech/produtosinternos/Paginas/Ferramenta%20SOAPP/O%20Meu%20Top%202020.aspx

Figura A.13: Template de Exemplo 2 (Pág.4)



Manual de Utilizador – Top 20

Finantech – Sistemas de Informação SA
Av. Fernão de Magalhães, 1862, 15º andar
4350-158 Porto Portugal
Telefone: +351 22 507 20 30
Fax: +351 22 507 20 59
Email: business@finantech.pt

Figura A.14: Manual de Exemplo 2 (Pág.1)

Finantech

Revisões ao Documento

| Data | Razão/Comentário | Versão | Versão | Responsável |
|------------|------------------|--------|--------------|-------------|
| 18-16-2015 | Top 20 - Manual | 1.0 | Renata Silva | Pedro Silva |

Figura A.15: Manual de Exemplo 2 (Pág.2)

Finantech

Índice

Índice.....3









O Meu Top 204

Manual de Utilizador

3 de 4

Figura A.16: Manual de Exemplo 2 (Pág.3)

O Meu Top 20

| O Meu Top 20 | | |
|-----------------|---------------------|---|
| Código Processo | Processo | |
| PRDDF-0073-068 | Gestão Documental |   |
| PRD-0071-070 | SSC - Screening |   |
| A&P-0002-002 | Implementação Soapp |   |
| AUS-0001-001 | Ausência |   |

À semelhança de outras listas, os registos devolvidos no "O Meu Top 20" têm a possibilidade de tomar determinado processo como favorito (primeira coluna de ícone) ou então retirar o processo da lista de favoritos (primeira coluna de ícone) e ainda inserir um registo de tempo diretamente em qualquer uma das linhas .

Esta lista é, automaticamente, atualizada todos os dias pelas 07 horas.

Figura A.17: Manual de Exemplo 2 (Pág.4)